protoDUNE online computing:
brief status report

Maxim Potekhin (BNL)

DUNE-LI Meeting

09/07/2016

Overview

- Events
- Documents
- Technology choices
- Plans

M Potekhin | protoDUNE Computing 9/7/2016
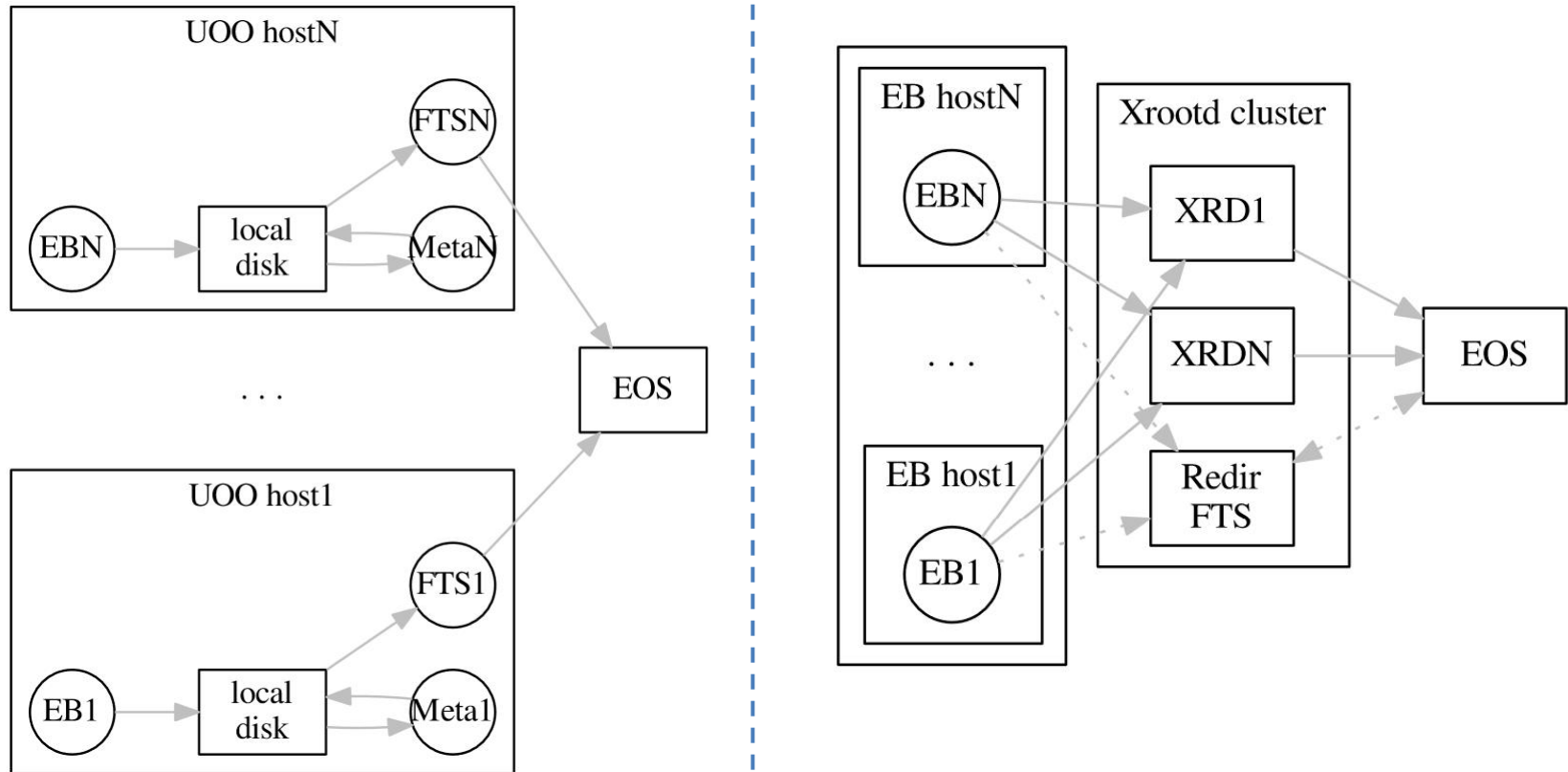
Events (recent and near future)

- A workshop was held at FNAL on July 28th-29th to discuss Workload Management
  - round table of experts from a few experiments and FIFE
  - provided definitions and better understood the requirements for DUNE
  - this is quite useful since it reflects a consensus (including FNAL), and the requirements better reflect the experience of a few experiments
  - the technology choice is still TBD
- Meeting with DAQ, artdaq and FTS experts at FNAL on August 23rd covering
  - use of high-speed storage (SSD) in the design of Event Builders in protoDUNE
  - methods for production of metadata and checksums
  - FTS interface(s) with online storage
- Coll. Meeting at FNAL Sept 12-15th
  - DAQ/online parallel sessions: will discuss design options
- CHEP2016 in October - a major HEP computing conference, we have an oral presentation on protoDUNE
- DAQ review at CERN on Nov 3-4/meeting with CERN-resident personnel
- XRootD workshop in mid-November 2016 - participation depends on whether XRootD remains the technology choice for the online buffer

BROOKHAVEN
NATIONAL LABORATORY

DUNE

Documents

- New design document: "The clustered storage option for the protoDUNE NP04 Online Buffer", DocDB 1628 - finalized in mid-August by Brett and Maxim
  - describes the proposal to re-purpose parts of the "neut" cluster at CERN to serve as the online buffer (admin: N.Benekos)
- "Minimum Multiplicity Requirements for ProtoDUNE DAQ", DocDB 1656 by Brett - a useful quantitative evaluation of the online data flow with a view to determine the minimum number of components required to support the projected data rates
- Technical content added to the DUNE Wiki at BNL:
  - https://dune.bnl.gov/wiki/XRootD_Buffer and other pages
- protoDUNE TDR - provisional deadline of Sept 1st
  - updated online computing sections, removed obsolete material, included new input
- DUNE Computing Model
  - requirements section updated (WMS)
  - protoDUNE section updated according to the current estimates of data characteristics

BROOKHAVEN
NATIONAL LABORATORY

DUNE

Technology choices for the online buffer

Two "old" design examples

BROOKHAVEN
NATIONAL LABORATORY

DUNE

Event Builders with attached SSD storage

- New design recently proposed by FNAL
- Introduces an additional layer of storage in the form of fast solid state disks attached to each Event Builder
- Has some advantages, but also creates more complexity as compared to fully networked storage e.g. xrootd
- Needs further justification and quantitative analysis, including costs
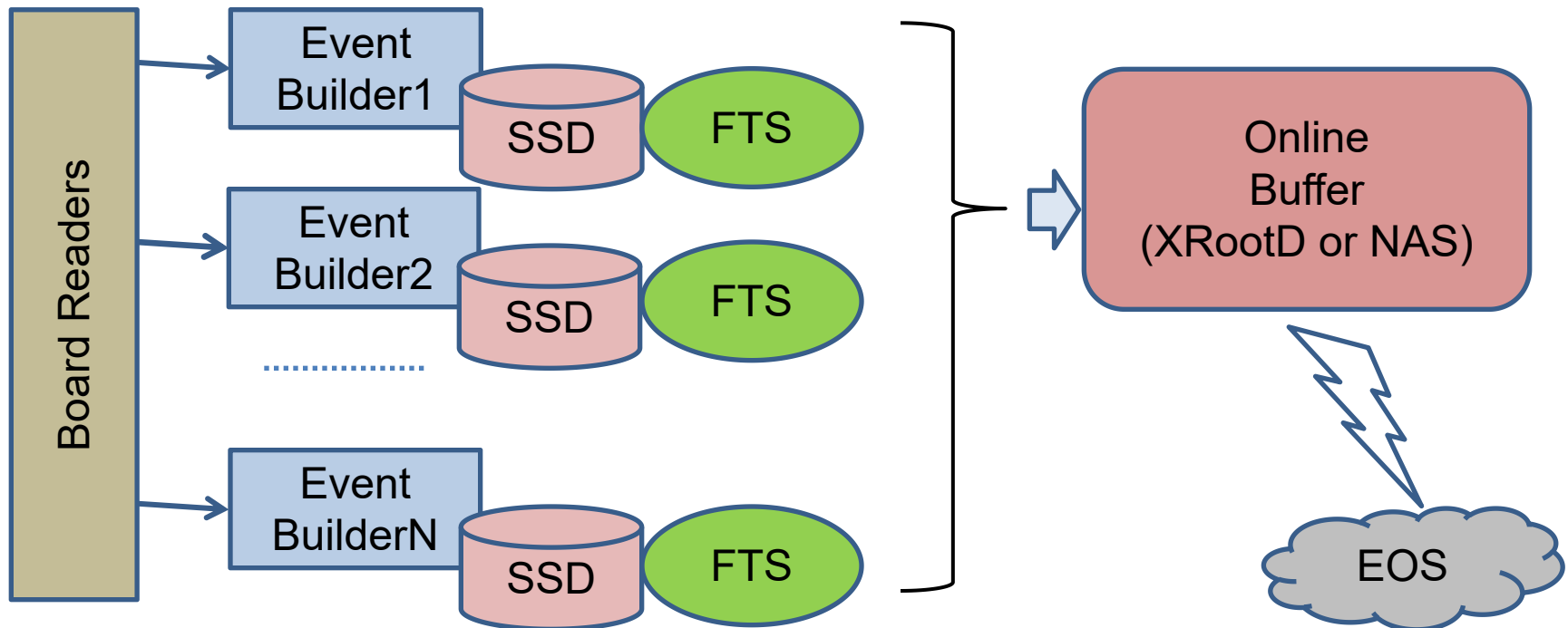
Technology Downselect

- Good progress has been made in understanding optimal ways to interface xrootd with F-FTS, with engagement of developers on both sides

- The buffer technology choice is currently in a state of flux due to new input from various parties involved (FNAL, CERN), but hopefully productive discussions will take place during the Coll. Meeting

- Type of design options being considered
  - "neut-based" xrootd storage cluster (the "original option")
  - xrootd cluster with new hardware purchase
  - Event Builders + SSD + xrootd storage cluster
  - Large NAS (industrial-grade network attached storage) + compute element

- Some of the crucial design decisions are:
  - where and how is the metadata formed?
  - what component is tasked with computing the checksum? how early in the pipe?
  - is the "3-day" buffer capacity requirement final, or it can be re-negotiated? TBD...
  - will protoDUNE be able to provide enough expertise for operating a NAS under extreme load?
  - how will the NAS scale and will it be cost-efficient?
  - what are the available power, space and cooling for the online buffer?

BROOKHAVEN
NATIONAL LABORATORY

DUNE

NAS+compute servers

- Built-in fault tolerance, configurable RAID

- Reasonable scalability (not entirely scale-out)

- Positive experience in ATLAS

- Cost may be significant because of sheer scale

- Not clear at this point how interface it with F-FTS

- Need to understand whether the proposed configuration has enoug CPU power (e.g. for checksums)

BROOKHAVEN
NATIONAL LABORATORY

DUNE

Why XRootD may be the best solution

- Good scaling-out behavior (there are examples of more than 600 nodes)
- Built-in fault tolerance
- Considerable expertise in the HEP/IF community
- Runs in user space and is fairly accessible for new practitioners
- Possibility to re-purpose existing hardware, and utilize it for other computational needs between the running periods
- Off-the-shelf components
- There are promising ideas about how to optimally interface it with F-FTS which is almost certain to become the protoDUNE data transport system
- Typical nodes come with healthy amounts of CPU power so the cluster can carry out some computations as required (e.g. checksums)

BROOKHAVEN
NATIONAL LABORATORY

DUNE

Plans

- We'll need to update our roadmap based on what's decided at FNAL in Sept.

- Right now the working assumption is that we are still on track to perform scalability test of the xrootd-based cluster at CERN

- The DUNE management wants to establish a closer connection between the plans for online systems between NP02 and NP04

- In my view, progress will depend on when the protoDUNE computing WG is established and staffed

- Outstanding items:

  - calibrations software: proposed algorithms need to be expressed in software, and will potentially be complex; could be time-critical

  - prompt processing: requirements need to be formulated, and technology platform chosen accordingly; needs effort

M Potekhin | protoDUNE Computing 9/7/2016

BROOKHAVEN
NATIONAL LABORATORY

DUNE